

Local Participants:

Julia, Costin, Eddie

Remote participants:

Pepe, Antonio, Alessandra

Meeting was recorded, recording can be found at:

<https://avprod.cern.ch/projects/clip/2650/>

Discussion during Julia's presentation*EGI portal data validation vs CMS accounting sources*

Antonio stressed again that CMS pilot implies complicated scheduling and in difference with other pilots its' consumption can not be compared with consumption of payloads

Julia replied, that yes, this is taken into account, however since currently there is no alternative information source centrally available for CMS, in case we can estimate average CMS pilot overhead, comparison of the payload usage in Dashboard and pilot usage in EGI portal can allow to detect problematic sites (problematic in terms of APEL accounting).

Julia mentioned investigation of the alternative solution for CMS accounting data being developed by Brian.

Antonio stated that CMS is working on a new central monitoring system which is needed by CMS itself and can be used for validation of the EGI portal, he is a part of this activity.

The conclusion from this discussion is that since Antonio takes part in the central monitoring activity in CMS he will follow up on the progress of the alternative data source for the EGI data validation and update the Task Force when we can try this alternative solution

EGI accounting reports for disk and tape usage

Julia asked Pepe, whether he thinks that any effort is currently required from the task force for validation of the disk and tape usage

Pepe replied that CMS does regular consistency checks (CMS catalogues and storage dumps) and though there is no regular checks against accounting portal numbers, in general they look fine.

He also mentioned that one should keep in mind that the space occupied by experiment data as experiment sees it, is less than space which is required at the site to handle this data properly, there is certain operational overhead, for example disk buffer in front of the tape storage. And it is necessary to take all space used into account and it all should be included in the pledge.

EGI Portal Testing and Validation

Pepe walked through the portal and shared the screen

First general statement:

Happy with the new version, much more professional from the looks and feel point of view and much faster

There were few comments to the entry page:

Units are mixed

When selecting different time range there is some confusion, looks like it changes time binning not time range

Then the main discussion was about the WLCG view

Main statement which everyone agreed upon:

We can not make plots or show percentage shares of the measurements which have different scale.

This is the case for non-normalized CPU. When CPU measurements which are not using the same reference CPU are shown in the same plot, the plot shows completely wrong shares, while these plots could be afterwards used by the funding agencies or scrutiny group members. Since EGI portal is an official tool we should make sure that things exposed there are fully validated and correct. Everyone agreed.

Julia asked what would Pepe suggest: to completely remove CPU tables, plots from the page, or to leave at least the table, but remove percentage shares and put a warning, that these measurements can not be compared between different instances.

Pepe agreed that leaving the table with a warning could be an option, but might be better to remove it completely.

We need to confirm with all people in the task force what they think, whether it would be still useful to expose CPU in the table but with a warning, or remove it completely.

Everyone agreed that the *main metrics exposed on the portal should be 'raw wall clock time' (hours) and 'work' (HS06)*. This again brings us back to the discussion, can we change the reporting mechanism and the whole processing chain in a way that for every job raw wall clock and work are reported and further processed on individual job records base, rather than transformation of the wall clock to work is performed using some 'per cluster' factors.

Pepe mentioned a problem of possible non-synchronization of the CPU reference with the measurements. If the site changes the attributes published in BDII which are used for time-to-work transformation, the transformation is done taking the latest published attributes in the account, while the measurements in the reports could be still done based on a different reference unit. This problem would be solved if time-work transformation happens at the node itself in the runtime and is reported with every job. MJF allows to enable it.

Another important feedback from Pepe. In the plot for CPU efficiency the averages in the rows and columns are not calculated correctly, it should be a waited average (all CPU divided by all wall clock), rather than average between all instances in the row or in the column.

In the cloud view, monetary cost is not clear what it is and should be removed.

The 'Reports' button

Pepe also looked into this one and concluded that it provides wrong data. Wall clock is not scaled by number of cores, etc...

Julia told, that the discussion about who and how to generate the WLCG accounting reports has to be held with Ivan. Whether WLCG can delegate generation of T1 and T2 reports to the EGI portal considering the fact that WLCG needs them to be the same as generated by REBUS. Before we decide on it, there is no point to validate what is published on the portal under 'reports'.

Julia told that she was going to clarify the situation with the reports with Ivan and might ask also Eddie to take part in this discussion since the code developed for REBUS might be easily reused by the the EGI portal as well.

Pepe volunteered to validate the reports when they are ready for validation

Julia's comments to the WLCG view

Julia suggested that the default plot located below the table in the WLCG view is stacked bar rather than line plot, but that there is a possibility to change it to line plot on user request.

Everyone agreed that we need a plot which shows wallclock consumption vs pledges, the one which Simone included in his input for the scrutiny group reports

Pepe showed also a plot which could be useful as well, which can show comparison of a particular instance vs average of other instances of this category

Julia asked Pepe to upload to the twiki page the file with other distributions he prepares for his site, so that people look through them and can be inspired by other ideas of useful plots. Pepe agreed.

Antonio also commented that at all plots all axes have to be labeled properly, both in terms of metric names and units

Presentation from Costin regarding ALICE accounting plans

The main idea of ALICE proposal is to run LHCb benchmark which apparently scales well with the ALICE workflows in the beginning of the pilot and based on it to calculate work out of wallclock in the runtime. In addition, it is suggested to send benchmark results along with the host identifier to the central DB, which would allow to accumulate history and to benchmark a particular node in a sort of dynamic way out of the running payloads. Costin demonstrated that normally the deviation of the measured scalability factors for a particular node is not substantial, which proves that such DB could be a good source for accounting 'time to work' transformation. Would be good to have such DB on the WLCG global scope.

During the following discussion it was mentioned that LHCb is also going to use the same scenario as ALICE and apparently ATLAS is interested as well.

The approach suggested by Pepe to use MJF was also discussed. However, Costin mentioned that according to ALICE tests, the LHCb benchmark scales much better with ALICE workflows compared to MJF and ALICE approach is more dynamic, which might be a better way when VMs are used.

Julia asked whether LHCb benchmark was tried with the CMS workflows. Antonio and Pepe did not know.

Julia mentioned that benchmark discussion is more relevant to the benchmarking WG. Alessandra confirmed that yes, this work is followed up by the HEPIX WG.

However, benchmarking area is very much related to the accounting and might be useful to understand whether all 4 experiments can use the same benchmark and whether all of them can be interested in the model described by Costin. We need to follow up on this proposal.

There was also a discussion of how accurate we aim to perform normalization, how different results we would get if we use per job normalization, rather than applying some per-cluster(site) normalization factor published to BDII as it currently happens. Pepe volunteered to make some investigation at his site.

Next meeting

Julia asked whether Pepe, Alessandra and Miguel who agreed to work on normalization investigation, could make some progress during next 2 weeks and present their findings/ideas/proposals at the next meeting. Agreed.

Next meeting is in two weeks

Action list

- Antonio will follow up on the alternative (rather than CMS Dashboard) central information source which can be used for EGI portal data validation
- Check with all task force members whether they agree to remove non-normalized CPU tables and plots from the WLCG view completely, or leave at least a table without percentage shares with a warning
- Julia will update the twiki page with all suggestions for the new portal. Pepe will upload a file with plots which might be useful to consider to implement on the portal
- Julia will follow up with Ivan the possibility and time estimation for generation of the accounting reports for T1 and T2 by the EGI portal with the same content and format as it is currently done by REBUS. It is possible that REBUS code could be re-used. Eddie will guide Ivan if it is the case.
- Pepe will investigate using PIC statistics, how much different the normalized wallclock will be if instead of using per-cluster normalization factor the individual per job normalization is performed
- Check with all 4 experiments whether ALICE benchmarking-accounting scenario can be of interest for other experiments (Julia and others)

- At the next meeting Pepe, Alessandra and Miguel will present their findings/ideas/proposals of the data transformation happening at all levels and possible better way to do it as a middle-term plan.